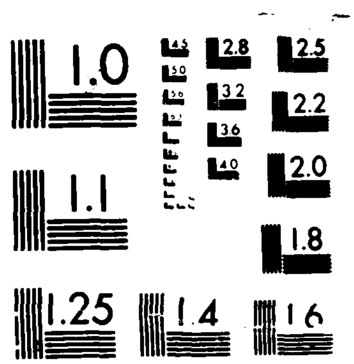
AD-A180 195 DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFEREN..(U) TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES A CHARNES ET AL. JAN 87 CCS-550 1/1

UNCLASSIFIED N00014-86-C-0398 F/G 12/4 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

Research Report CCS 558

DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFERENCE SETS

by

A. Charnes
W.W. Cooper
J. Rousseau
J. Semple

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

(12)

Research Report CCS 558

DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFERENCE SETS

by

A. Charnes
W.W. Cooper
J. Rousseau
J. Semple

January 1987

This research was partly supported by ONR Contracts N00014-86-C-0398 and N00014-82-K-0295, and National Science Foundation Grants SES-8408134 and SES-8520806 with the Center for Cybernetic Studies, The University of Texas at Austin. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director

College of Business Administration, 5.202
The University of Texas at Austin
Austin, Texas 78712-1177
(512) 471-1821

A

ABSTRACT

DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFERENCE SETS

By A. Charnes, W.W. Cooper, J. Rousseau and J. Semple

The University of Texas at Austin

Serious mathematical and computational errors and misstatements culminating in erroneous characterizations of Data Envelopment Analysis (DEA) models and methods and their relationship to the "axiomatic" production models of Shepard type by Fare, Hunsaker and others are corrected together with new exposition and contrast of current DEA methodology with the Shepard axiomatic modelling types. The stochastic base of DEA is shown to be uncertainty, not risk. A new computationally effective "extended additive" (EA) model is developed to handle processes with input thresholds and output ceilings and thereby not subsumable under Shepard axiomatics.

KEY WORDS

Data Envelopment Analysis

CCR Ratio Model

Multi-objective Programming

Efficiency Analysis

Research Report CCS 558

DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFERENCE SETS

By A. Charnes, W.W. Cooper, J. Rousseau and J. Semple

Introduction

The Management Science paper (Vol. 32, 2, February 1986) of Färe and Hunsaker, "Notions of Efficiency and Their Reference Sets," contains serious mathematical and computational errors and misstatements culminating in erroneous characterizations of Data Envelopment Analysis (DEA) models and methods and their relationship to the "axiomatic" production models of Shepard type. In particular, every (sic) DEA example solution presented is erroneous, plus the stated ratio model itself is erroneous despite their bibliographic citations of papers [1], [2] containing the correct non-Archimedean model. Again, the paper presents axioms stated in obsolete mathematical terminology plus incomplete specification so that equivalence claims for models and more appear to be technically unsupported.

In the following, we correct such errors, exposit and contrast the DEA methodology with that which rests on the complex Shepard axiomatic production theory. The cited examples of the latter structures, re efficiency considerations, are subsumed in those CCR ratio models which have "reference," or "production possibility," sets of conical type. These conical sets, as recognized by Shepard and Farrell [26], do not accurately describe production possibilities for many processes. We therefore bring to the fore some of these real phenomena and data aspects and their implications for desiderata in reference sets, modeling, computation, informational presentation and management evaluation. In particular we present a new computationally effective extension of the (non-ratio C^2GS^2) DEA model of [3], which we call the "extended additive" (EA) model, which can properly take account of processes with input thresholds and output ceilings and thereby a whole range of discretionary to non-discretionary variable configurations not subsumable under the Shepard axiomatics.

Non-Archimedean Errors

As presented orally and in print (xerox) at a session of the Rutgers University Productivity Conference in 1981 of Dr. A. Dogramaci and attended by R. Färe and as presented in the Färe - Hunsaker (F-H) bibliography [1] [2], the non-Archimedean CCR ratio model is (in simpler notation)

$$
\begin{array}{lll}
 & \max h = u^T y_0 / v^T x_0 & \\
(1) & u^T y_j / v^T x_j \leq 1, & j = 1 \ldots, n \\
 & -u^T / v^T x_0 \leq -\epsilon e^T & \\
 & -v^T / v^T x_0 \leq -\epsilon e^T &
\end{array}
$$

where (x_j, y_j) is the observed input - output vector of DMU_j, 0 ("zero") designates the DMU whose efficiency is being estimated, e^T is a vector of "ones" and ϵ is the (positive) non-Archimedean infinitesimal.

By the Charnes - Cooper transformation $\mu = tu$, $\upsilon = tv$, $tv^T x_0 = 1$, it transforms into the linear programming problem (R) with dual (DEA), the Data Envelopment side,

$$
\begin{array}{llcl}
 & \text{(R)} & & \text{(DEA)} \\
 & \max h = \mu^T y_0 & & \min \theta \quad -\epsilon e^T s^+ - \epsilon e^T s^- \\
 & \upsilon^T x_0 = 1 & & \\
(1.1) & \mu^T Y - \upsilon^T X \leq 0 & & Y\lambda - s^+ = y_0 \\
 & -\mu^T \leq -\epsilon e^T & & \theta x_0 - X\lambda - s^- = 0 \\
 & -\upsilon^T \leq -\epsilon e^T & & \\
 & & & \lambda, s^+, s^- \geq 0
\end{array}
$$

Where $Y \overset{\Delta}{=} [y_1, \ldots, y_n]$, $X \overset{\Delta}{=} [x_1, \ldots, x_n]$

Evidently the optimal value will be of "complex number" form $a - b\epsilon$, with $a = \theta^*$ representing the "scale" efficiency and b representing $e^T s^{*+} + e^T s^{*-}$, the sum of optimal slacks (input surpluses and output shortfalls).

Instead, F - H <u>erroneously</u> render (1) as

<u>ERRONEOUS MODEL</u>

$$\max h = u^T y_0 / v^T x_0$$

(2)
$$u^T y_j / v^T x_j \leq 1, \quad j = 1, \ldots, n$$
$$-u^T \leq -\epsilon e^T$$
$$-v^T \leq -\epsilon e^T$$

Using subscript 2 in place of the confusing subscript 0 of F-H, their example problems may be written as

(2.1)
$$\max \quad u/(2''v_1 + 5''v_2)$$
$$u/(2''v_1 + 5''v_2) \leq 1$$
$$u/(2''v_1 + 6''v_2) \leq 1$$
$$u, v_1, v_2, \geq \epsilon$$

and

(2.2)
$$\max \quad u/(2''v_1 + 6''v_2)$$
$$u/(2''v_1 + 5''v_2) \leq 1$$
$$u/(2''v_1 + 6''v_2) \leq 1$$
$$u, v_1, v_2 \geq \epsilon$$

where the quotation marks denote multiplication by a (the same) power of 10.

Setting $w_1 = u/(2''v_1 + 5''v_2)$, $w_2 = u/(2''v_1 + 6''v_2)$, since the u and v_i are non-negative, we have <u>always</u> $w_1 \geq w_2$, thus the second constraint is redundant. For optimality in both (2.1) and (2.2) we have $w_1^* = 1$ or $u^* = 2''v_1^* + 5''v_2^*$

Thereby $w_2^* = (2''v_1^* + 5''v_2^*) / (2''v_1^* + 6''v_2^*) = 1 - 1''v_2^*/(2''v_1^* + 6''v_2^*)$.

Evidently DMU_1 is efficient ($w_1^* = 1$) in all of these examples and

(2.1.1) $$v_1^* = 1,\ v_2^* = 1,\ u^* = 2'' + 5''$$

is one of the infinite number of optimal solutions, which further, need not involve the non-Archimedean infinitesimal !

Because of $w_2 = 1 - 1''v_2/(2''v_1 + 6''v_2)$ for optimality in (2.2), there is no maximum but only a supremum value of one ("efficiency"). For, taking $v_1 = pv_2$ and $v_2 = 1$, we have

(2.2.1) $$h = w_2 = 1 - 1/(2p + 6) \to 1 \text{ as } p \to \infty$$

Even further, as may be suspected from the above, (2), which are F - H's equations (2.2) on page 238, may be reduced to an equivalent form without the non-Archimedean infinitesimal ϵ ! Setting $\bar{u} = \epsilon u$, $\bar{v} = \epsilon v$, (2) becomes

(3) $$\begin{array}{lll} \max & \bar{u}^T y_0/\bar{v}^T x_0 & \\ & \bar{u}^T y_j/\bar{v}^T x_j & \leq 1, \quad j = 1, \ldots, n \\ & \bar{u}^T & \geq e^T \\ & \quad \bar{v}^T & \geq e^T \end{array}$$

The F - H "solution", $h = 1/(1 + \epsilon)$, $v_2 = \epsilon$, $v_1 = (1 - 5\epsilon)/2$, $u = 1$ on p. 239, equation (2.4) and its accompanying statement "Since the solution to problem (2.3) is not a real number, the value of h cannot be computed" are erroneous, witness (2.2.1) and (3).

The F - H "solutions" of their example 2, $h = 1/8$, $u = 1$, $v_1 = v_2 = 10^{-6}$ and $h = 1/7$, $u = 1$, $v_1 = v_2 = 10^{-6}$ (for, respectively, DMU_2 and DMU_1) and their accompanying statement "Thus no one DMU is efficient" are erroneous since by (2.2.1) $\sup(h) = 1$ for DMU_2 and by (2.1.1), $h = 1$ for DMU_1.

Using instead on example 2 the correct CCR ratio model (1) and the DEA side of the linear programming equivalent (1.1), from which all Data Envelopment Analysis computations have been performed heretofore, we obtain by trivial observation the optimal basic solutions

(1.3.1) $\theta^* = 1,\ \lambda_1{}^* = 1,\ s_2{}^{*-} = 0$ with dual optimum $\mu^* = 1,\ \nu_1{}^* = (1 - 5''\epsilon)/2$
$\nu_2{}^* = \epsilon$ with $h^* = \mu^* = 1$ for DMU_1.

and

(1.3.2) $\theta^* = 1,\ \lambda_1{}^* = 1,\ s_2{}^{*-} = 1''$ with dual optimum $\mu^* = 1 - 1''\epsilon$,
$\nu_1{}^* = (1 - 1''\epsilon)/2''$, $\nu_2{}^* = \epsilon$ with $h^* = \mu^* = 1 - 1''\epsilon$ for DMU_2.

Thus, correctly, DMU_1 is rated efficient and DMU_2 is not.

For the last few years, a real efficiency value, if desired, has been determined from the optimal real $\theta, \lambda, s^+, s^-$ as

(1.3.3) $$\theta^*[1 - (e^T s^{*+} + e^T s^{*-}) / (e^T x_0 + e^T y_0 + e^T s^{*+} + e^T s^{*-})]$$

In example 2, DMU_1 thus has value 1 and DMU_2 has value approximately 0.9.

Most vitally, the CCR ratio model always has, via the dual theorem of linear programming applied to (1.1), a basic optimal solution pair, hence a true minimum for DEA and a true maximum for (R). Evidently, from their own example, the erroneous F - H model does not.

Computation and Informatics

Computation for (1.1) has been available from the first (1951) in Charnes' non-Archimedean simplex method and solution of the "degeneracy" problem [4], in classes for the past 35 years and in the 1961 Charnes - Cooper text, "Management Models and Industrial Applications of Linear Programming" [5], by the use of additional columns and rows containing the (base field) coefficients of the non-Archimedean quantities. The NONARC code of Dr. Iqbal Ali (of The University of Texas at Austin) will also solve such non-Archimedean linear programming problems by a 2-phase process similar to the Phase I, Phase II process of usual simplex method codes, which actually solves a similar non-Archimedean problem. Additionally, ordinary simplex codes can be used for ϵ sufficiently small numerically relative to the other matrix data reciprocals. The infinitesimal must however be taken larger than the numerical zero tolerance. Routinely also, before computation, linear programming codes scale the matrix of coefficients by change of units or multiplication across rows so that the entries lie in the 0 - 100 range.

Toward both purposes as mentioned in [2], cited in the F - H bibliography, the coefficient of θ in the functional of (1.1) has been enlarged to 100 (corresponding to percent scale efficiency) or more. And one can even multiply the functional by ϵ^{-1} leaving θ with the only non-Archimedean coefficient. As shown by the exact solutions for u^*, v_1^*, x_2^* in (1.3.1) and (1.3.2), numerical difficulties are certain for computation in example 2 with quotation marks corresponding to 10^6 and $\epsilon = 10^{-6}$. Enlargement of the coefficient of θ in the functional and/or scaling as stated in [2] has sufficed to remove such difficulties in all past applications handled by professionals in mathematical programming computation.

The major problem in DEA analysis computations has not been and is not such numerical niceties, but rather in solving the informatics problem of providing the solution to DEA problems for all the DMUs with preservation and extraction of the detailed information needed for analysis. For example, one requires "window analyses" based on efficiency values [6], and "envelopment maps" for assessing stability, possible data errors and robustness of the empirical Pareto-Koopmans optimal production function together with time series behavior [7]

This is in addition to providing solution detail (on the DEA side) for every DMU. Further, contrary to the impression given by the F-H paper, there exists not one but a multiplicity of DEA models, the software for which, thus far, has been elided so far as possible so that with a single software package one can call up a DEA analysis for any one of these types.

Development of such efficient packages has been possible because at most only the right hand side, a column and the functional need be changed in passing from one DMU to another. Also, the invariant multiplicative DEA model [8] has the same formal structure as the "additive" C^2GS^2 model [3]. Because the analyses required from most models demand extensive computation, it is vital when developing new models for other production possibility sets to tie in their specifications as closely as possible to the computational capabilities of existing software modules and the additional data which can be accessed easily through simple extensions.

Shepard Axiomatics

Shepard formulated a theory of production [12] for a restricted class of production technologies which culminated in an elegant transform theory between production and cost aspects called Shepard-Samuelson Duality. More recently others such as his student R. Färe and the economist R. Russell, have focused on extending this theory to encompass efficiency considerations which thereby involve certain abstractly defined frontier sets related to the Shepardian characterization of efficiency. Their work involved similarly the complex mathematical constructs of point-to-set maps ("correspondences," in obsolete terminology) of Shepard so that to exercise this theory on particular technologies it is necessary to identify them as satisfying Shepard-type axioms.

These axioms typically refer to a point-to-set map L transforming an output vector u into the set L(u) of all input vectors x which can produce the output vector u. A key axiom is

L4: L(u) is a closed correspondence

This axiom is not verified in the F-H paper. Indeed, "closed correspondence" is not even defined there or in other papers in the field. The Färe, Grosskopf, Lovell monograph [9] (p 203) gives an incomplete (and typographically erroneous) definition. Russell, perhaps confused by the archaic terminology and absent topologies he attempts to correct, renders this axiom incorrectly as

L4'. L(u) is a closed set (of input points!).

To define a closed map requires the specification of three topologies --- one for the domain, one for the range, and one for the product topology of the two. A map may be closed or not depending on which topologies are prescribed for these three topological spaces. A most clear exposition and examples are given on p. 312-313 of the Narici and Beckenstein text [10]. The correct definition follows.

Let S and T denote topological spaces and S X T their topological product space. Let f be a map from S into T. The graph of f, G(f), is the set of pairs {(s, f(s)): s ∈ S}.

Definition: S is a "closed map" iff G(f) is a closed subset of S X T in the product topology.

We note that the definition of graph on p. 25 of Färe, Grosskopf, and Lovell, equation (2.1.3), is thereby incorrect since it involves pairs of input output vectors rather than pairs $(u, L(u))$ of output vector u and image $L(u)$, the latter a point in the space of all <u>sets</u> of input vectors x. No topology is specified for the latter as points nor for the topological product. Thus the Färe, et al, identifications are invalid since the axioms to be satisfied are inadequately or erroneously defined.

Data Envelopment Analysis

Contrary to the impression given in the Färe and Hunsaker paper, Data Envelopment Analysis is a methodology which in relation to notions of efficiency and reference sets has developed important sets of models and new relations between classical efficiency notions, economic production theory, as well as new auditing and operational aspects of management of productive processes undertaken by a multiplicity of organizational or response units. The first DEA model, the non-Archimedean CCR ratio model, established a connection between the classic scientific-engineering notion of efficiency and the Farrell economic production theory notion by generalizing the scientific-engineering efficiency definition to multiple input, multiple output systems and to relative efficiency. These were connected by mathematical programming duality as in the (R) and (DEA) of equations (1.1).

The combination of multiple inputs (multiple outputs) into a single virtual input (virtual output) is possible in an infinite number of other ways than in (1.1), e.g., a multiplicative one by raising inputs (outputs) to nonnegative powers and multiplying them together to obtain a single virtual input (virtual output). The resulting ratio model, on taking logarithms of input and output entries, reduces to a linear programming model with, once again, the dual DEA side corresponding to a measure of relative efficiency for an explicitly (analytically) stated production possibility set [11]. The form of this which gives an efficiency measure independent of the units in inputs and outputs yields dually always a piece-wise Cobb-Douglas economic production function [8].

The key paper of Charnes, Cooper, Golany, Seiford and Stutz [3] established the fact that the extremal models of Data Envelopment Analysis, on the DEA side, were the Charnes-Cooper test for Pareto-optimality of a test point (input-output vector) [13], [14]. Thus, the essence of the DEA methodology is in providing a test for multi-objective optimality of a specified point (here a possible DMU input-output vector) chosen from a specified reference set (here production possibility set) which yields simultaneously quantitative values of input surpluses and output shortfalls needed to obtain efficiency.

Recognizing this connection, Charnes, Cooper, and Thrall [15] developed an Archimedean characterization of efficiency types for the CCR ratio model in which scale efficiency and technical efficiency are chacterized by the dimensionality of the set of optimal solution vectors to the correspondent of (R) in (1). Via semi-infinite programming this has been generalized to an infinite number of DMU's by Charnes, Cooper, and Wei [16] with corresponding generalizations of other models forthcoming. Further, another direction of generalization of the ratio models has been to arbitrary closed convex cones for the virtual multipliers and for the production possibility sets and data envelopments by Charnes, Cooper, Wei and Huang [17] with extensions to other models forthcoming as well. The dimensionality construct in Charnes, Cooper and Thrall [15] has been simplified (and generalized) to interiority properties relative to the cone of virtual multipliers [16]. Thus, this extension and the semi-infinite extension represent generalizations to possibly nonpolyhedral reference and virtual multiplier sets which involve thereby nonlinear programming and duality relations.

Before proceeding to a new computationally-effective DEA model, we need to correct a misimpression by the noted economist, P. Schmidt [18], to the effect that Data Envelopment Analysis is a deterministic and not a statistical methodology. The fact is that all the DEA methodology from Farrell onward involve determinations of relative efficiency and of efficient (frontier) production functions based on sample observations. By the mathematical definition of the word "statistic" the quantities calculated are all statistics. What is true is that so far we have only a bare beginning of statistical theory for these statistics [3]. A particular problem is how one should characterize "waste" (input surpluses) and "possible achievement failures" (output shortfalls) stochastically. This is not addressed adequately by parametric formula fitting according to maximum likelihood or other classical statistical estimation schemes which, incidentally, have proven substantially inferior to DEA methods in correctly estimating known frontier formulae from random samples [19]. Failing such a theory, devices such as window analysis [6] and envelopment maps [7] have been developed to expand sample sizes to achieve

methods in correctly estimating known frontier formulae from random samples [19]. Failing such a theory, devices such as window analysis [6] and envelopment maps [7] have been developed to expand sample sizes and to achieve robustness in efficiency characterizations and production function determinations as well as to analyze time series or dynamic effects.

The Extended Additive (EA) Model

Meaningful determination of relative efficiency should involve comparisons of a DMU to actual production possibilities only. In some applications, the model employed has efficiency determination and measurement of input surpluses and output shortfalls referred to efficient points which are not possible. Such relative efficiency (and effectiveness) determinations are thereby invalid. Many of these applications have nevertheless proved worthwhile because the most important determinations made have been (correctly) with reference to efficient inputs and outputs that were possible. Examples of this kind have occurred when some input variables are nondiscretionary, i.e., are not controllable by a DMU manager, such as temperature or weather or unemployment rate. The modification suggested by Banker and Morey [20] involving a change in the functional is not appropriate since Pareto-optimal (efficient) DMU's may fail to be recognized because the modified functional is incorrect for the Charnes-Cooper test plus the partly conical character of the reference set (production possibility set) may lead to an impossible efficient referent. Again, many production processes involve input thresholds below which the process does not work and output ceilings above which output is impossible (as, for example, in sales in a market, the number of purchasing households in the area cannot exceed the number of households in that area.).

In the key paper [3] "Foundations of Data Envelopment Analysis for Pareto Optimal Empirical Production Functions," a model which today we call the "additive" model was put forward which remedied the production possibility set difficulty of cones and provided a basis for other more valid reference sets. We now extend this model to one which also takes care of thresholds and ceilings as well as permitting a variable to range from discretionary to nondiscretionary in a way which is trivially implementable on existing DEA software.

The Charnes-Cooper test for Pareto Optimality (here Pareto-Koopmans Optimality) of DMU_0, using as reference set the convex hull of the input-output vectors of DMU_1 to DMU_n, may be written (see [3]) as

$$\begin{aligned} \min \quad & -e^T s^+ - e^T s^- \\ & Y\lambda - s^+ = y_0 \\ \text{(A.1)} \qquad & -X\lambda - s^- = -x_0 \\ & e^T\lambda = 1 \\ & \lambda, s^+, s^- \geq 0 \end{aligned}$$

If one wishes this value to be independent of the units of measurement we alter the functional to

$$\text{(A.2)} \qquad -e^T D^{-1}(y_0)s^+ - e^T D^{-1}(x_0)s^- \equiv -d^w(s^+, s^-)$$

where $D(y_0)$, $D(x_0)$ are diagonal matrices with the y_0 or x_0 component entries. (If some components of y_0 or x_0 are zero we use the unique Moore - Penrose <u>generalized</u> inverse which has zeros instead of reciprocals for the zero components.) For <u>either</u> functional we obtain <u>the same</u> DMUs as efficient since

$$\text{(A.2.1)} \qquad \alpha\,(e^T s^+ + e^T s^-) \leq e^T D^{-1}(y_0)s^+ + e^T D^{-1}(x_0)s^- \leq \beta\,(e^T s^+ + e^T s^-)$$

where α, β are respectively the minimum and maximum of the non-zero entries of both $D^{-1}(y_0)$ and $D^{-1}(x_0)$ and therefore $\alpha, \beta > 0$. It is useful to have the (A.1) functional form both for theory and practice since the efficiency hyperplane it defines $(e^T s^+ + e^T s^- = 0)$ does not depend on any particular x_0, y_0.

When DMU_0 is inefficient, its <u>precise</u> efficiency measure is not the important consideration, but rather the input surpluses and output shortfalls needed for efficiency (i.e., to be on the economic (Pareto-Koopmans Optimal) production function surface) together with the "facet" of efficient referent DMUs determining the local production function. These are the

significant managerially meaningful quantities. Using the (A.2) functional, a computationally easy efficiency score using the weighted l_1 -metric of (A.2) is

$$\text{(A.2.2)} \quad \exp\{-d^w(s^+, s^-)/[d^w(y_0, x_0) + d^w(s^+, s^-)]\} \approx 1-d^w(s^+, s^-)/[s+m + d^w(s^+, s^-)]$$

(The latter may be interpreted as a ratio of distance measures as in [21]). Note that this units invariant score may be applied equally well to the multiplier of θ^* in (1.3.3).

We now present the Extended Additive (EA) model for efficiency rating of DMU_0 as.

$$\text{(EA)} \quad \begin{array}{llll} \min & -e^T s^+ - e^T s^- \;, \quad \text{or} \quad -e^T D^{-1}(y_0)s^+ - e^T D^{-1}(x_0)s^- & & \\ & Y\lambda \; -s^+ \qquad\quad = \; y_0 & & \\ & -X\lambda \qquad\quad -s^- = -x_0 & & \\ & e^T\lambda \qquad\qquad\quad = \; 1 & & \\ & \qquad\qquad s_i^- \le \beta_i x_{0i}, & i = 1, \ldots, m & \\ & \qquad s_r^+ \qquad \le \gamma_r y_{0r}, & r = 1, \ldots, s & \\ & \qquad\qquad\qquad \lambda, s^+, s^- \ge 0 & & \end{array}$$

In this, by adjusting the value of β_i from zero to one, we go from a nondiscretionary input to a completely discretionary input. Practically, since the numerical value of a nondiscretionary input is not that precise, one may use always a small value of β_i dependent on the character of the input for DMU_0. Raising the value of β_i corresponds to not allowing the referent input value to go below a specified level less than the ith input value of DMU_0. I.e., the referent input must exceed this threshold input level.

Similarly, by adjusting the value of γ_r from zero to one we go from the fixed output r level of DMU_0 to anything permitted by the other constraints. Raising the value of γ_r corresponds to not allowing the referent output value to exceed a specified level above the rth output level of DMU_0. I.e., the referent output cannot exceed this ceiling output level. The use of

such individual upper bounds on the slacks poses no computational problem in Data Envelopment Analysis since every linear programming code of any merit today incorporates the individual upper bound algorithmically without requiring its explicit specification as a constraint.

Thus the EA model permits the automatic incorporation of real local restrictions on many processes which are not permitted by the other existent models. As such, the associated Pareto-Koopmans efficiency surface is a more valid representative of the true nature of the processes.

Conclusions

To summarize, contrary to the Färe, Hunsaker, et al, perjoratives, Data Envelopment Analysis is a statistical methodology whose models consist of the Charnes-Cooper test for a Pareto-Koopmans minimum applied to test point observations from a specified reference set. The model may be a linear programming model, or a transformably convex programming model in which the functional may or may not be linear, and in which the reference set may or may not be polyhedral and may (as in EA) or may not vary with the test point (=DMU).

The CCR, Multiplicative, BCC [22], Additive and EA models provide different production possibility (reference) sets for Data Envelopment Analysis. Its connection with multi-criteria programming is through the Charnes-Cooper test for multi-criteria optimality, here Pareto-Koopmans Optimality, which test is one of the few available constructive means for determining such optima [23], e.g. efficient DMUs.

Data Envelopment Analysis determines "facets" of similar efficient DMUs in estimating the unknown economic production function in the neighborhood of *each DMU with respect to the* overall referent set. The input surpluses and output shortfalls of each DMU that DEA determines, rather than the precise numerical score assigned to inefficiency, are the central valuable results. Statistically, or stochastically, DEA deals with estimation from samples in situations of uncertainty (wherein the joint distribution of the random variables, input surpluses and output shortfalls, are unknown) rather than classical situations of risk (wherein the joint distribution is known). The stochastic models of Aigner, Schmidt and others have dealt only with risk. The difference this makes in difficulty and lack of statistical theory may be appreciated from the work of Dvoretsky, Kiefer and Wolfowitz (1952) on the much simpler Inventory Problem [24], [25] with known function. Thus, statistical theory development is an important challenge for DEA research.

REFERENCES

[1] Charnes, A., W.W. Cooper, A.Y. Lewin, R.C. Morey, and J. Rousseau, "An Approach to Positivity and Stability Analysis in DEA," The University of Texas at Austin, Center for Cybernetic Studies, Report CCS 434, 1982, and Annals of Operations Research, 2, pp 139-156, 1985.

[2] Charnes, A. and W.W. Cooper, "The non-Archimedean CCR Ratio for Efficiency Analysis: A Rejoinder to Boyd and Färe," European Journal of Operations Research, 15, pp 333-334, 1984.

[3] Charnes, A., W.W. Cooper, B. Golany and L. Seiford, "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions," Journal of Econometrics, 30, pp. 91-107, 1985, an extension of Charnes, A., W.W. Cooper, and L. Seiford, "An Empirical DEA Production Function," The University of Texas at Austin, Center for Cybernetic Studies CCS Report 396, April 1981.

[4] Charnes, A., "Optimality and Degeneracy in Linear Programming," Econometrica, April, 1952.

[5] Charnes, A. and W.W. Cooper, Management Models and Industrial Applications of Linear Programming, vols. 1 and 2, J. Wiley and Sons, 1961.

[6] Charnes, A. C.T. Clark, W.W. Cooper and B. Golany, "A Developmental Study of Data Envelopment Analysis in Measuring the Efficiency of Maintenance Units in the U.S. Air Force," Annals of Operations Research, 2, pp. 95-112, 1985

[7] Charnes, A., W.W. Cooper, N. Eechambadi and F. Phillips, "Competition and Dynamics in Data Envelopment Analysis," presented in N. Eechambadi's PhD. Dissertation, The University of Texas at Austin, Efficiency Analysis of Market Response and the Marketing Mix. Extending Data Envelopment Analysis to a Competitive Environment, 1985

[8] Charnes, A., W.W. Cooper, L. Seiford, and J. Stutz, "A Units Invariant Multiplicative Measure in Efficiency Analysis," Operations Research Letters, 1983

[9] Färe, R., S. Grosskopf and C.A.K. Lovell, The Measurement of Productive Efficiency, Kluwer-Nighoff Publishing, 1985

[10] Narici, L. and E. Beckenstein, Topological Vector Spaces, Marcel Dekker Inc., 1985

[11] Charnes, A., W.W. Cooper, L. Seiford and J. Stutz, "A Multiplicative Model for Efficiency Analysis," Socio-Economic Planning Sciences, v. 6, pp. 223-224, 1982.

[12] Shepard, R.W., Cost and Production Functions, Princeton University Press, 1950.

[13] Ben-Israel, A., A. Ben-Tal and A. Charnes, "Necessary and Sufficient Conditions for a Pareto-optimum in Convex Programming," Econometrica, 45, pp. 811-820, 1977.

[14] Ben-Israel, A., A. Ben-Tal and S. Zlobec, Optimality in Nonlinear Programming: A Feasible Directions Approach, (pate 89), J. Wiley and Sons, 1981.

[15] Charnes, A., W.W. Cooper and R. Thrall, "Classifying and Characterizing Efficiencies and Inefficiencies in Data Envelopment Analysis," Operations Research Letters, 5, 3, pp. 105-109, 1986.

[16] Charnes, A., W.W. Cooper and Q. Wei, "A Semi-infinite Multi-criteria Programming Approach to Data Envelopment Analysis with Infinitely Many Decision-making Units," The University of Texas at Austin, Center for Cybernetic Studies Report CCS 551, September 1986.

[17] Charnes, A., W.W. Cooper, Q. Wei and Z. Huang, "Cone Ratio Data Envelopment Analysis and Mult-objective Programming," The University of Texas at Austin, Center for Cybernetic Studies Report CCS 559, December 1986.

[18] Aigner, D.J., C.A.K. Lovell and P. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics, v. 6, pp. 21-37, 1977.

[19] Banker, R., A. Charnes, W.W. Cooper and A. Maindiratta, "A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology," Working Paper 85-08-2, IC^2 Institute, The University of Texas at Austin, 1985, and in Application of Modern Production Theory: Efficiency and Productivity, A. Dogramaci and R. Fare (Ed.), 1986.

[20] Banker, R. and R. Morey, "Efficiency Analysis for Exogenously Fixed Inputs and Outputs," submitted October 1983 to Operations Research.

[21] Charnes, A., W.W. Cooper, B. Golany, L. Seiford and J. Stutz, "A Dimensionless Efficiency Measure for Departures from Pareto Optimality," The University of Texas at Austin Center for Cybernetic Studies Report CCS 480, February 1984.

[22] Banker, R., A. Charnes and W.W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," Management Science, 30, 9, pp 1078-1092, 1984.

[23] Zlobec, S. "Two Characterizations of Pareto Minima in Convex Multicriteria Optimization, Aplikace Matematiky, v. 29, pp. 342-349, 1985.

[24] Drovetsky, A., J. Kiefer and J. Wolfowitz, "The Inventory Problem, I," Econometrica, July 1952.

[25] Drovetsky, A., J. Keifer and J. Wolfowitz, "The Inventory Problem, II," Econometrica, July 1952.

[26] Charnes, A., W.W. Cooper and A. Schinnar, "Transforms and Approximations in Cost and Production Function Relations," Omega, 10, 2, pp. 207-211, 1982.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: CCS 558	2. GOVT ACCESSION NO.: AD-A180 195
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): DATA ENVELOPMENT ANALYSIS AND AXIOMATIC NOTIONS OF EFFICIENCY AND REFERENCE SETS	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER: CCS 558
7. AUTHOR(s): A. Charnes, W.W. Cooper, J. Rousseau, J. Semple	8. CONTRACT OR GRANT NUMBER(s): N00014-86-C-0398, N00014-82-K-8295
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Cybernetic Studies, UT Austin, Austin, TX 78712	10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research (Code 434), Washington	12. REPORT DATE: January 1987
	13. NUMBER OF PAGES: 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Data Envelopment Analysis, CCR Ratio Model, Multi-objective Programming, Efficiency Analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Serious mathematical and computational errors and misstatements culminating in erroneous characterizations of Data Envelopment Analysis (DEA) models and methods and their relationship to the "axiomatic" production models of Shepard type by Fare, Hunsaker and others are corrected together with new exposition and contrast of current DEA methodology with the Shepard axiomatic modelling types. The stochastic base of DEA is shown to be uncertainty, not risk. A new computationally effective "extended additive" (EA) model is developed to handle processes with input thresholds and output ceilings and thereby not subsumable under Shepard axiomatics.

END
6-87
DTIC